Analyses of Subgenomic Promoters of *Hibiscus Chlorotic Ringspot Virus* and Demonstration of 5′ Untranslated Region and 3′-Terminal Sequences Functioning as Subgenomic Promoters

Weimin Li[1] and Sek-Man Wong[1,2]*

Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543,[1] and Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604[2]

Hibiscus chlorotic ringspot virus (HCRSV), which belongs to the genus *Carmovirus*, generates two 3′-coterminal subgenomic RNAs (sgRNAs) of 1.4 kb and 1.7 kb. Transcription start sites of the two sgRNAs were identified at nucleotides (nt) 2178 and 2438, respectively. The full promoter of sgRNA1, a 118-base sequence, is localized between positions +6 and −112 relative to its transcription start site (+1). Similarly, a 132-base sequence, from +6 to −126, defines the sgRNA2 promoter. Computer analysis revealed that both sgRNA promoters share a similar two-stem-loop (SL1 + SL2) structure, immediately upstream of the transcription start site. Mutational analysis of the primary sequence and secondary structures showed further similarities between the two subgenomic promoters. The basal portion of SL2, encompassing the transcription start site, was essential for transcription activity in each promoter, while SL1 and the upper portion of SL2 played a role in transcription enhancement. Both the 5′ untranslated region (UTR) and the last 87 nt at the 3′ UTR of HCRSV genomic RNA are likely to be the putative genomic plus-strand and minus-strand promoters, respectively. They function well as individual sgRNA promoters to produce ectopic subgenomic RNAs in vivo but not to the same levels of the actual sgRNA promoters. This suggests that HCRSV sgRNA promoters share common features with the promoters for genomic plus-strand and minus-strand RNA synthesis. To our knowledge, this is the first demonstration that both the 5′ UTR and part of the 3′ UTR can be duplicated and function as sgRNA promoters within a single viral genome.

The production of subgenomic RNAs (sgRNAs) is one of the strategies used by positive-stranded RNA virus for expression of 3′-proximal genes from their polycistronic genome. The sgRNAs, which are coterminal with the genomic RNA (gRNA), are generated by three possible mechanisms, including internal initiation on a minus-strand template, premature termination, and discontinuous transcription (17, 26). Viruses such as *Brome mosaic virus* (BMV) and *Turnip crinkle virus* (TCV) adopt the strategy of internal initiation to generate sgRNAs (19, 24).

The sgRNA promoter, an internal sequence in the minus strand of gRNA, plays an important role in sgRNA synthesis. During this process, the sgRNA promoter sequence is recognized by the RNA-dependent RNA polymerase. So far, sgRNA promoters have been mapped and characterized in several viruses, and the size ranges from about 20 nucleotides (nt) (9, 10, 15) to over 100 nt (5, 13, 21). Most of these promoters are mapped to sequences upstream of the transcription start site. However, the sgRNA promoter of *Beet necrotic yellow vein virus* (3) and those of sgRNA2 and sgRNA3 of *Barley yellow dwarf virus* (BYDV) (13) are located primarily downstream.

Viruses with more than one sgRNA often contain only short stretches of sequence homology around the transcription start site. The sgRNA promoter sequences often form a stem-loop

(SL) structure, which is proposed to facilitate the interaction with the transcriptase. Generally, the combination of primary sequence and secondary structural elements is required for sgRNA transcription in vivo (5, 12, 20, 25). The transcription start site is indispensable for promoter activity among some viruses including BMV, *Tobacco mosaic virus* (TMV), BYDV, *Potato virus X*, and *Alfalfa mosaic virus* (AMV). Alteration or deletion of this nucleotide greatly diminishes or eliminates promoter activity (2, 5, 11, 13, 21).

Hibiscus chlorotic ringspot virus (HCRSV), a member of the genus *Carmovirus* in the family of *Tombusviridae*, has a plus-strand RNA genome of 3,911 nucleotides containing seven open reading frames (ORFs). Two 3′-coterminal sgRNAs, sgRNA1 (1.47 kb) and sgRNA2 (1.73 kb), are generated during infection and serve as mRNAs to express two movement proteins (MPs) and coat protein (CP). Their transcription start sites were mapped to nt 2177 and 2437, respectively (8) (Fig. 1A).

In this study, we aimed to map the promoter regions of sgRNA1 and sgRNA2 and to analyze the function of the primary sequence and putative secondary structure for promoter activity. In addition, we tested if the 5′ untranslated region (UTR) or the last 87 nt at the 3′ UTR of the HCRSV genome were able to function as individual sgRNA promoters in vivo.

MATERIALS AND METHODS

Plasmid constructions. All plasmids were originally constructed from plasmid p223, a full-length cDNA clone of HCRSV (8). To confirm the transcription start site of sgRNA, mutants sg1A/U and sg1G/C were created with forward mutagenic primers F-S1-A/T and F-S1-G/C and a reverse primer, R-Hpa2555. PCR products were doubly digested with NcoI and HpaI and ligated into p223. Mutants of sg2C/G and sg2G/C for sgRNA2 promoter were made by overlapping

* Corresponding author. Mailing address: Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore. Phone: 65-65162976. Fax: 65-67792486. E-mail: dbswsm@nus.edu.sg.
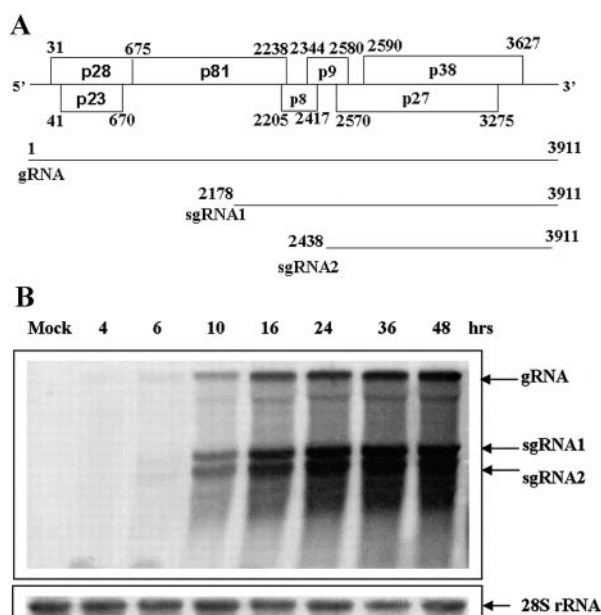
A

B

FIG. 1. Time course of accumulation of HCRSV sgRNA1 and sgRNA2 in kenaf protoplasts. (A) Schematic diagram of the HCRSV genome. The open boxes represent ORFs with encoded protein products. Three single horizontal lines represent gRNA, sgRNA1, and sgRNA2, respectively. Numbers indicate the nucleotide positions on the HCRSV genome. (B) Northern blot analysis of total RNA from kenaf protoplasts transfected with wild-type HCRSV in vitro transcripts. The protoplasts were harvested at the indicated time intervals postinoculation. Bands corresponding to HCRSV gRNA and the sgRNA1 and sgRNA2 are shown with arrows. The level of 28S rRNA was shown to indicate the relative amount of samples loaded.

PCR. In the first round of PCR amplification, two regions of p223 were amplified with primers F-Nco2131 and R-S2-G/C and primers F-S2-G/C and R-Hpa2555 (for mutant sg2C/G) or primers F-Nco2131 and R-S2-C/G and primers F-S2-C/G and R-Hpa2555 (for mutant sg2G/C). The two intermediate products of the first round were mixed and used as templates for the second round of PCR for amplification with F-Nco2131 and R-Hpa2555. The resulting PCR product was digested with NcoI and HpaI and cloned into p223.

To map the sgRNA1 promoter, a mutant, p223BS, was generated by overlapping PCR to introduce two restriction sites, BglII and SphI, into the cDNA of HCRSV at position 3126. Two series of deletion mutants were constructed as follow. For 5′ border mapping, fragments containing deletions were duplicated by using a set of forward primers (spanning nt 2080 to 2099, 2106 to 2125, 2127 to 2146, 2145 to 2164, 2160 to 2179, and 2175 to 2194, with an additional BglII site at the 5′ end) and a reverse primer, SG1-Sph(−)2183, flanking a SphI site. For 3′ border mapping, corresponding fragments were generated using a forward primer, Sg1-Bgl(+)2052, with a BglII site and a set of reverse primers complementary to nt 2212 to 2231, 2181 to 2200, 2169 to 2188, 2164 to 2183, 2161 to 2180, and 2157 to 2176, with a SphI site at the end. The resulting PCR products were digested with BglII and SphI and inserted into p223BS.

To define the 5′ border of the sgRNA2 promoter, the nucleotide A at position 2240 was replaced with a T to create a new BglII restriction site, resulting in mutant p223B. Deletion fragments were amplified by a set of forward primers (spanning nt 2312 to 2331, 2346 to 2365, 2376 to 2395, 2399 to 2418, 2418 to 2437, and 2433 to 2453, with an additional BglII site at the 5′ end) and the reverse primer R-Hpa2555 and then treated with BglII and HpaI and ligated into p223B. The 3′ border of sgRNA2 was mapped using the same strategy, with progressively deleted sequences generated by a forward primer corresponding to nt 2555 to 2577 (containing an HpaI site at position 2563) and a set of reverse primers (spanning nt 2508 to 2527, 2453 to 2472, 2437 to 2456, 2424 to 2443, 2421 to 2440, 2419 to 2438, and 2417 to 2436 with an HpaI site at the 5′ end).

To analyze the secondary structures of sgRNA1 and sgRNA2 promoters, mutants with duplicated promoters were constructed by PCR amplification of the putative promoter regions with the primers listed in Table 1, containing flanking

BglII or SphI restriction sites, and by cloning the BglII- and SphI-digested PCR fragments into p223BS. By using the same strategy as described above, four additional mutants, pmini-sg1 and pmini-sg2, p5UTR, and p87, were constructed with the corresponding primers (Table 1), respectively.

Northern blot analysis and secondary structure prediction. One microgram of p223 was linearized with SmaI, and infectious RNA was produced by in vitro transcription (MEGAscriptT7; Ambion). Kenaf (*Hibiscus cannabinus* L.) protoplasts were prepared and transfected with 10 μg of RNA according to the method described in a previous study (16). Total RNA (5 μg) extracted from inoculated protoplasts at 24 h postinoculation (hpi) was analyzed by Northern blot hybridization (23). A probe complementary to nt 3126 to 3911 of the HCRSV genome was prepared with a DIG labeling mix (Roche) and used for detection of gRNA and sgRNA accumulation. The promoter activity was quantified as the ratio of sgRNA intensity to that of the gRNA as described previously (12). All mutants were evaluated in two to four independent experiments. The RNA MFOLD version 2.3 from the website www.bioinfo.rpi.edu/applications /mfold (27) was used to predict the secondary structure of RNA sequence.

RESULTS

Accumulation of HCRSV subgenomic RNAs in kenaf protoplasts. The accumulation of both sgRNA1 and sgRNA2 as well as gRNA in kenaf protoplasts inoculated with wild-type (wt) transcripts of a full-length HCRSV cDNA clone, p223, was examined from 4 to 48 hpi by Northern blot analysis using a 3′ HCRSV-specific probe. The levels of the two sgRNAs and the plus-strand gRNA were observed to increase steadily from 6 to 48 hpi (Fig. 1B).

Modification of the transcription start sites of HCRSV subgenomic RNAs. Mutagenesis of the transcription start site of sgRNA generated via internal initiation can greatly decrease and even abolish sgRNA synthesis (2, 5, 11, 13, 21). Therefore, we changed the putative transcription start site of nucleotide A at 2177 to U for sgRNA1 and the nucleotide C at 2437 to G for sgRNA2 to create mutants sg1A/U and sg2C/G, respectively (Fig. 2A). Northern blot analysis of total RNA from transfected protoplasts showed that the corresponding sgRNAs of the two mutants still accumulated to detectable levels. However, mutation of G at 2178 to C (sg1G/C) and mutation of G at 2438 to C (sg2G/C) completely abolished accumulation of sgRNA1 and sgRNA2, respectively (Fig. 2B).

Previously, the transcription start sites of sgRNA1 and -2 were mapped by primer extension to positions 2177 and 2437, respectively (8). However, in this study we discovered that the actual transcription start sites are located at positions 2178 and 2438, respectively. The errors in assignment of positions in the previous report were due to miscalculations of nucleotides.

Mapping of the boundaries of the sgRNA1 promoter. The initiation nucleotide of sgRNA1 promoter is mapped to nt 2178, which is within the ORF of the 81-kDa protein that is part of the putative RNA-dependent RNA polymerase complex. It means that the sgRNA1 promoter overlaps with the complementary strand of the p81 ORF. To avoid interruption of virus replication, the following strategy was employed to map this promoter. Three fragments, ranging from nt 1982 to 2312, nt 1982 to 2262, and nt 2052 to 2262, containing the putative sgRNA1 promoter, were duplicated and inserted into p223BS at position 3126 in the CP ORF (Fig. 3A). The resulting constructs each produced an ectopic sgRNA (sgRNA3), indicating that full-length sgRNA1 promoter was contained within the complementary sequence of nt 2052 to 2262.

Subsequently, a series of smaller fragments of different sizes were amplified by PCR and introduced into p223BS (Fig. 4A).

TABLE 1. Primers used for mapping of sgRNA1 and sgRNA2 of HCRSV[a]

Expt	Primer name	Primer sequence (5'–3')
Modification of sgRNA transcription start site	F-S1-A/T	CTCCATGGACATTGAGCTTGAGTTTGCCCCGG*T*^{2147}GGGA
	F-S1-G/C	CTCCATGGACATTGAGCTTGAGTTTGCCCCGGAC^{2148}GGA
	F-S2-C/G	GACTTGAGAACCCTC*G*^{2437}GGGAAAAATTGCT
	R-S2-C/G	AGCAATTTTCCCC*G*^{2437}GAGGGTTCTCAAGTC
	F-S2-G/C	GACTTGAGAACCCTCCC^{2438}GGAAAAATTGCT
	R-S2-G/C	AGCAATTTTCCCGG^{2438}GAGGGTTCTCAAGTC
	R-Hpa2555	CTTTCCAGTGTTAACTACTATAT
	F-Nco2131	GAGGAGTTTTATGACTCCATGG
Structural analysis of sgRNA1 promoter	SGI-Bgl(+)2067	GTAGATCTAGAAGAAAGCCAGGGTGAGT
	SGI-Sph(−)2183	CAGCACATGCTTTCCCTCCGGGGCAAACTC
	SGI-Bgl(+)2052	GTAGATCTTTCATGAACAAACAGAAG
	SGI-SL2A1(+)	CAAGATCTACTTCAGAAGAAGCCAGGGTGAGT
	SGI-SL2A2(−)	CAAGATCTCAGAAGAAGCCAGGGTGAGTTTCTTCAAAG^{2006}*GGA*^{2008}TTGGGGGTGGACCCTATGG
	SGI-SL2B(+)	CAAGATCTTTC^{2179}*GCA*^{2177}CCGGGGCAAACTCAAGCT
	SGI-SL2B(−)	CAAGATCTCAGAAGAAGCCAGGGTGAGTTTCTTCAAAG^{2006}*GCA*^{2008}TTGGGGGTGGACCCTATGG
	SGI-SL2C1(+)	CAAGATCTGCTTTCCCTCC^{2174}*CCCC*^{2171}CAAACTCAAAGCTTT^{2101}*CCCC*^{2104}TGGACCCTATGGCTCAAA
	SGI-SL2C2(−)	CAAGATCTGCTTTCCCTCC^{2174}*CCCC*^{2171}CAAACTCAAAGCT
	SGI-SL2D1(+)	CAAGATCTGCTTTCCCTCCGGGGCAAACTCAAGCTTT^{2101}*CCCC*^{2104}TGGACCCTATGGCTCAAA
	SGI-SL2D1(−)	CAAGATCTGCTTTCCCTCCGGGGCAAACTCAAGCTCAATGT
	SGI-SL2D2(−)	CAAGATCTGCTTTCCCTCCGGGGCAAAC^{2151}Δ^{2165}ATGGAGTCATAAAACTCC
	SGI-SL2E1(−)	CAGCATGCTTTCCCTCCGGGGCAAACTCAAGCTCAATGTCCATGG^{2143}Δ^{2138}AGTACTCCTCAATGGCCATTT
	SGI-SL2E2(−)	CAGCATGCTTTCCCTCCGGGGCAAACTCAATGTCCATGGAGT^{2143}*GTATTT*^{2138}ACTCCTCAATGGCCATTT
Structural analysis of sgRNA2 promoter	SG2-Bgl2312(+)	CAAGATCTACTTCAAGCAGCAGTGCACAAAGC
	SG2-Sph2443(−)	CTGCATGCTTTCCCGGAGGGGTTCTCAAG
	SG2-SL2A(+)	CTAGATCTACTTCAACAAGCAGTGCACAAAGCTAATAAGCACGATGAAGCTG^{2353}*AAAG*^{2356}TGGGGGGTTCTTTCGTTTT
	SG2-SL2A(−)	CTAGATGC^{2443}*AA*^{2341}CCCGGAGGGGTTCTCAAGT
	SG2-SL2B(+)	CTAGATCTACTTCAAGCAGTGCACAAAGCTAATAAGCACGATGAAGCTGTTTCTGG^{2360}*CCCAAGA*^{2366}TTCGTTTTTTGTTTGCTGATA
	SG2-SL2B(−)	CTAGATGC^{2434}*CCCAAGA*^{2428}CAAGTCCACGTCAAAGT
	SG2-SL2C(+)	CTAGATCTACTTCAACAAGCAGTGCACAAAGCTAATAAGCACGATGAAGCTGTTTCTGGGGGGGTTCT^{2427}Δ^{2405}TTAAATGGTAAACCTCAACTT
	SG2-SL2C(−)	CTAGATCTACTTCAACAAGCAGTGCACAAAGCTAATAAGCACGATGAAGCTGTTTCTGGGGGGGTTCT^{2427}Δ^{2376}TTGCTGATAAAAGTTGAGG
	SG2-SL2D(+)	CTGCATGCTTTCCCGGAGGGGTTCTCAAGTCCACGTCAAAGTTAATGTAACCTCAACTT
	SG2-SL2D(−)	CTGCATGCTTTCCCGGAGGGGTTCTCAAGTCCACGTCAAAGTTAATGTTAATGGTAA^{1999}*CTATT*^{1995}CCTCAACT^{1986}*AATGG*^{1982}TGCAACAAAAA CGAAAGAAC
	SG2-SL2E1(−)	CTGCATGCTTTCCCGGAGGGGTTCTCAAGTCCACGTCAAAGTTAATGTTAATGGTAA^{1987}*GGAGTTGA*^{1994}TTATCAGCAACAACAAAAGA
	SG2-SL2E2(−)	CTGCATGCTTTCCCGGAGGGGTTCTCAAGTCCACGTCAAAGTTAATGTTAATGGTAA^{1987}*GGAGTTGA*^{1994}TTATCAGCAACAACAAAAGA
Defining minimal sequence of promoters	Mini-SG1-B	CTAGATCTCCTTTGGGGGTGGATTTG
	Mini-SG1-S	CTAGATGCTTTCCCTCCGGGGCAAATCCACCCCAAAGG
	Mini-SG2-B	CTAGATCTGGGAAACCGTGGGCAAGTT
	Mini-SG2-S	CTGCATGCTTTCCCGAGGGTTCTAGAACCCCCAGAAA
Duplication of the 5' UTR and last 87 nt of 3' UTR	87-Bgl	CTAGATCTGGGCTGCCTCACAACTATG
	87-Sph	CTGCATGCTGGAATCCATAAAAATCCTT
	5UTR-Bgl	CTAGATCTGGGAAACCGTGGGCAAGTT
	5UTR-Sph	CTGCATGCAGAGAATTCGGAAACTTGC

[a] Altered bases are in italics; deletions are shown as Δ; and restriction recognition sequences of NcoI, HpaI, BglII, and SphI are underlined. The numbers in superscript indicate altered or deleted nucleotide positions in the HCRSV genome.
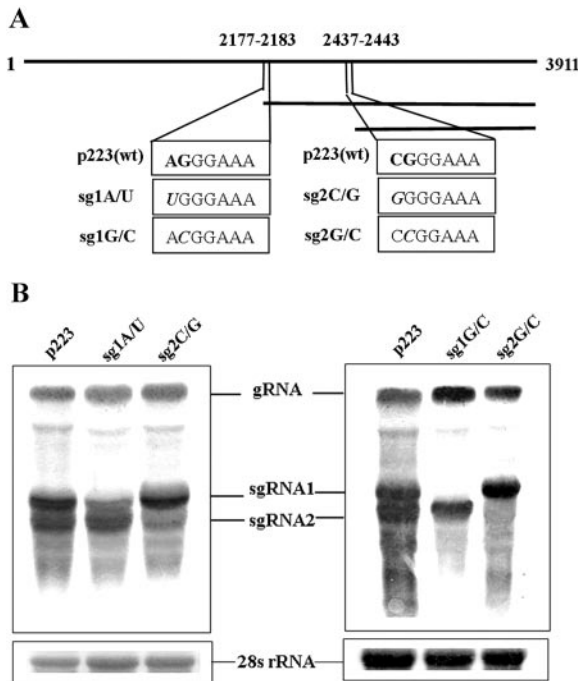
FIG. 2. Modifications of transcription start sites of HCRSV sgRNA1 and sgRNA2. (A) Mutagenesis of the putative transcription start sites for sgRNA1 and sgRNA2. The nucleotides are shown in the boxes with mutations in boldface and italics, and the mutant names are listed on the left of each box. (B) Northern blot analysis of total RNA from kenaf protoplasts transfected with p223 and mutant transcripts (24 hpi). Bands corresponding to gRNA and sgRNAs are indicated.
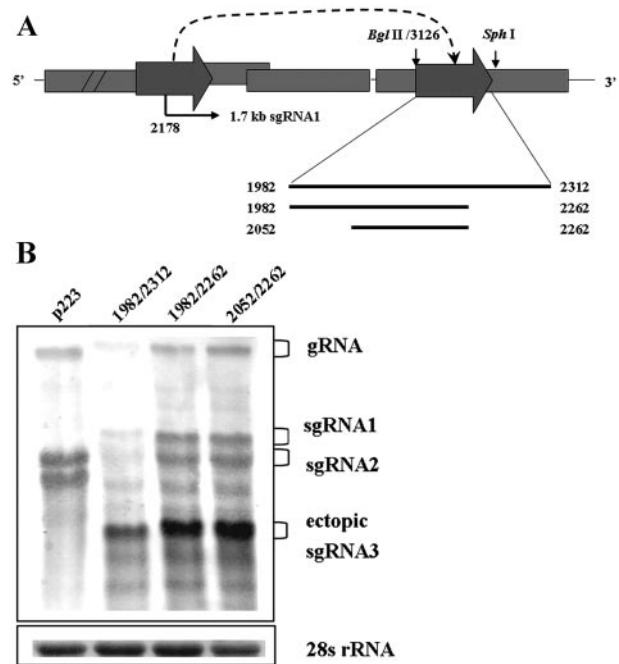

FIG. 3. Sequence requirements for ectopic expression of the sgRNA1 promoter. (A) Schematic diagram of constructs in which sequence encompassing the context of sgRNA1 transcription start site was duplicated and inserted into the BglII and SphI sites in the coat protein ORF of the p223BS mutant. Position 2178 is shown as the sgRNA1 transcription start site. Three horizontal lines with original position numbers denote the length of the duplicated sequences, which are indicated with filled arrowheads in the construct. (B) Northern blot analysis of sgRNA accumulation in protoplasts (24 hpi) transfected with in vitro transcripts of p223 and the three test constructs. Bands corresponding to the ectopic sgRNA3 are indicated.

In vitro transcripts from these constructs were used to transfect kenaf protoplasts, and the total RNAs were subjected to Northern blot analysis. To delineate the 5′ border of the sgRNA1 promoter, deletions were made from nt 2052 to 2174, upstream of the transcription start site (Fig. 4A). Compared to mutant 2052/2262 (Fig. 4B, lane 1), sgRNA3 accumulation by mutant 2080/2262 is strongly reduced (Fig. 4B, lane 2), whereas sgRNA3 accumulation by mutant 2106/2262 (Fig. 4B, lane 3) and other 5′ deletion mutants (Fig. 4B, lanes 4 to 7) is below the level of detection. Thus, the 5′ boundary of the sgRNA1 promoter was located between nt 2052 and 2079.

To map the 3′ border, stepwise deletions were carried out from positions 2262 to 2176. The mutants with different duplicated fragments from nt 2052 to 2183 could still produce a distinguishable sgRNA3 (Fig. 4C, lanes 1 to 5). However, the quantity of sgRNA3 generated by the construct with a duplicated fragment from nt 2052 to 2180 was greatly decreased (Fig. 4C, lane 6), and sgRNA3 was not detectable with the construct containing a duplicated fragment from nt 2052 to 2176 (Fig. 4C, lane 7). Therefore, the 3′ border of the full-length sgRNA1 promoter was mapped to position 2183 of the HCRSV gRNA, five bases downstream of the transcription start site. Taken together, the full length of the sgRNA1 promoter was defined as a fragment of 132 bases (complement of nt 2052 to 2183) containing 126 bases upstream and 5 bases downstream of the initiation site, respectively.

Mapping of the boundaries of sgRNA2 promoter. The initiation site of the sgRNA2 promoter is at position 2438. It

appears that sgRNA2 promoter overlaps with the putative MP ORF p9, which is dispensable for replication. Therefore, this promoter was tested in its original location (Fig. 5A).

To map the 5′ border of the promoter, we made a construct, p223B, with a BglII site at position 2237. A set of mutants was created with progressively larger deletions from nt 2237 to 2178 (Fig. 5A). Deletion of nt 2237 to 2311 had no effect on the accumulation of the sgRNAs (Fig. 5B, lane 1), while a mutant with a deletion from nt 2237 to 2345 was able to produce sgRNA2, but at a lower level (Fig. 5B, lane 2). No sgRNA2 was detected when the deletion was extended to position 2375 (Fig. 5B, lane 3).

To delineate the 3′ border of the sgRNA2 promoter, deletion from nt 2562 to the transcription start site at 2178 was performed. Surprisingly, deletions of nt 2562 to 2444 resulted in an increase, rather than a decrease, of sgRNA2 level (Fig. 5C, lanes 1 to 4). However, after continuing deletion of 3 nt from 2443 to 2441, the level of sgRNA2 was substantially reduced (Fig. 5C, lane 5). Constructs with deletions up to nt 2439 and 2437 eliminated sgRNA2 altogether (Fig. 5C, lanes 6 and 7), indicating that the 3′ border of the promoter was located between nt 2440 and 2443. These results indicated that the sgRNA2 promoter contained 132 bases (complement of nt 2312 to 2443), including 126 bases upstream and 5 bases downstream of the transcription start site, respectively.
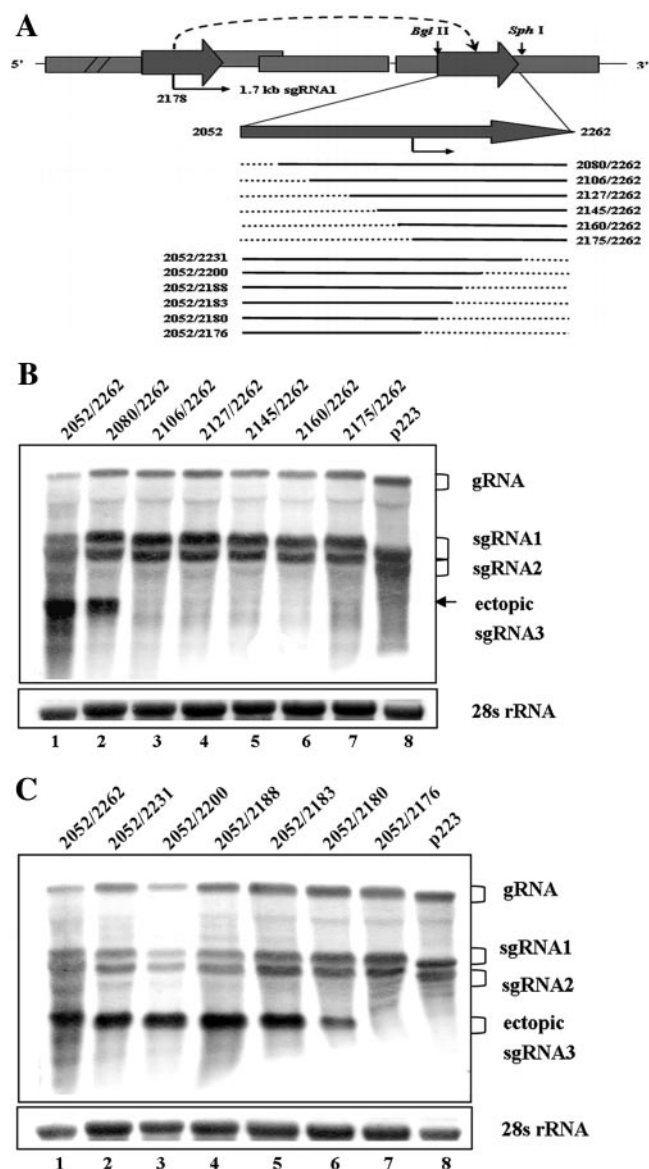
FIG. 4. Mapping of the boundary of sgRNA1 promoter by ectopic expression of sgRNA3. (A) Map of the constructs that contain duplicated copies of the sgRNA1 promoter (filled gray arrow) inserted into the BglII and SphI sites of p223BS. The transcription start site at position 2178 is shown by an arrow. The duplicated sequences are indicated with horizontal solid lines, along which the listed numbers are the original position numbers of 5′ and 3′ ends of the duplicated sequences. The dotted lines represent the deleted sequences. Each deletion mutant is named with the position numbers of two endpoints of the duplicated sequence. (B) Northern blot analysis of the 5′ border of sgRNA1 promoter in kenaf protoplasts. (C) Mapping of the 3′ border of sgRNA1 promoter in kenaf protoplasts. Bands corresponding to gRNA, sgRNA1, sgRNA2, and ectopic sgRNA3 are indicated.
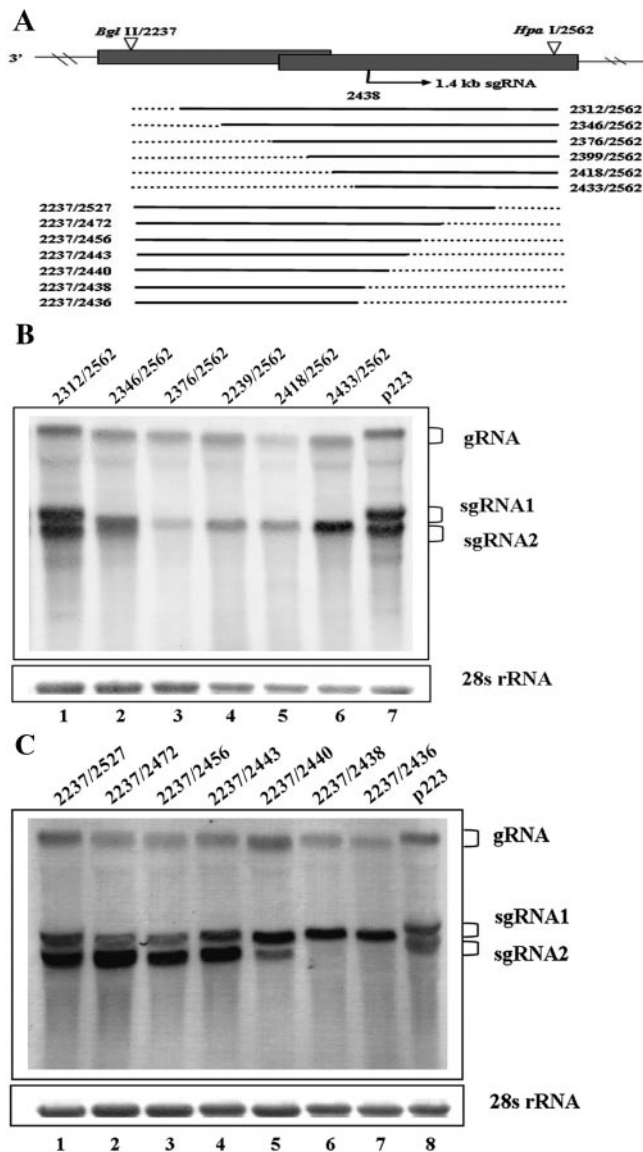
FIG. 5. Mapping of the boundary of sgRNA2 promoter. (A) Schematic representation of deletion mutants. The transcription start site at position 2438 is shown by an arrow. The dotted lines represent the deleted regions. The potential promoter region within each mutant is shown as a horizontal solid line. The numbers denote its 5′ and 3′ positions. (B) Northern blot analysis of the 5′ border of sgRNA2 promoter in kenaf protoplasts. (C) Mapping of the 3′ border of sgRNA2 promoter in kenaf protoplasts. Bands corresponding to gRNA, sgRNA1, and sgRNA2 are indicated.

The two subgenomic RNA promoters of HCRSV share similar secondary structures. RNA-MFOLD predicted possible secondary structures in the mapped sgRNA1 and sgRNA2 promoter sequences. Alignment of the nucleotide sequences of the two promoters showed only 36% identity. However, the two promoters appear to possess a similar secondary structure

consisting of two stem-loops (Fig. 6A and 7A). For the sgRNA1 promoter (Fig. 6A), 1SL1, the complementary region of nt 2068 to 2091, was located between −110 and −87 relative to the transcription start site (+1), and 1SL2 was positioned between −82 and +2, which was complementary to nt 2096 to 2179. For the sgRNA2 promoter (Fig. 7A), 2SL1, the complementary region of nt 2312 to 2350, was located between −126 and −88 relative to the transcription start site, and 2SL2, complementary to nt 2353 to 2443, was positioned between −85 and +6. Interestingly, both secondary structures can be
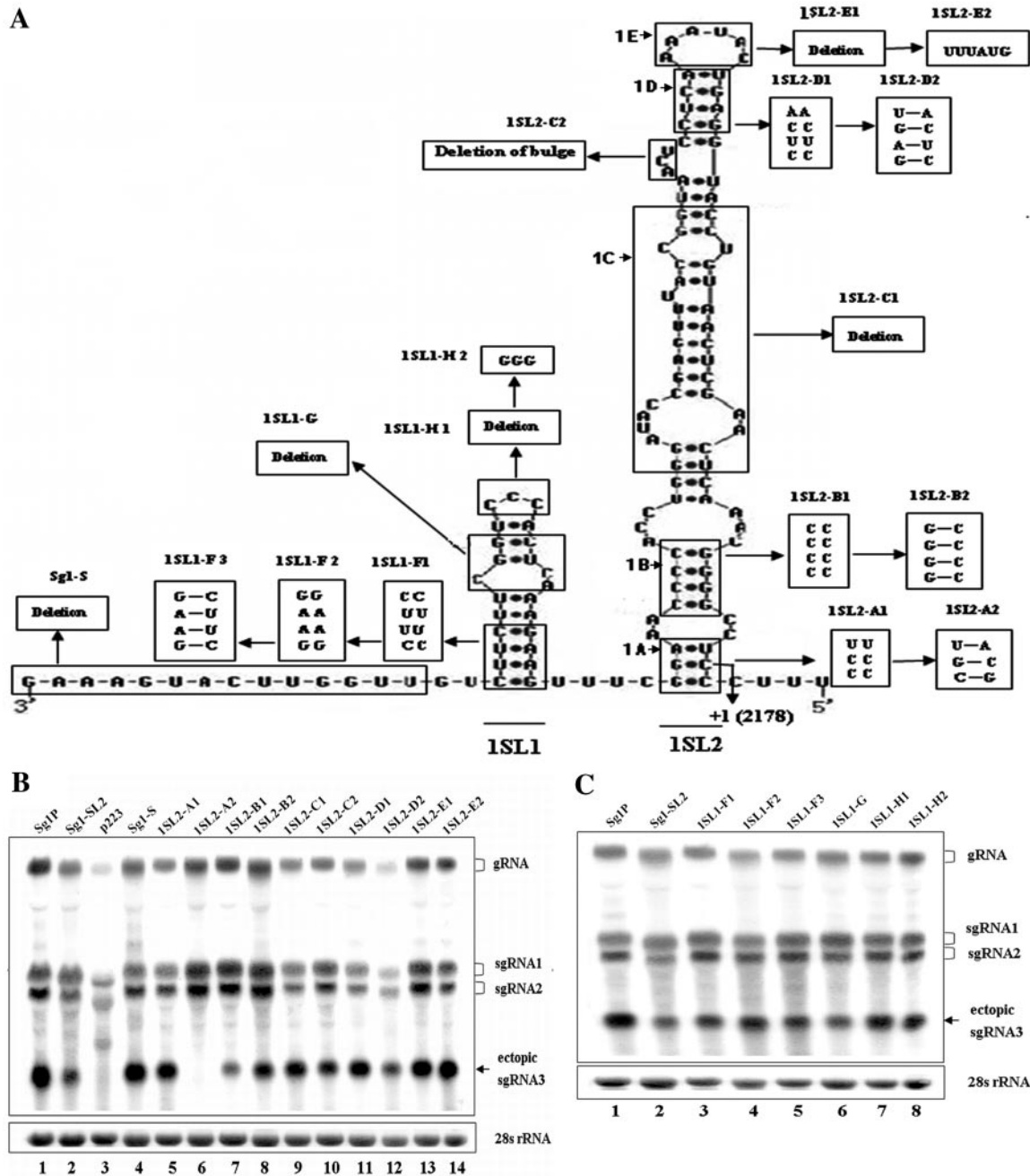
FIG. 6. Analyses of RNA sequence and secondary structure of sgRNA1 promoter. (A) Mutations in the regions of the duplicated sgRNA1 promoter. The putative two-stem-loop structure (1SL1 and 1SL2) was predicated by MFOLD at 25°C. Boxes contain nucleotides inserted as substitutions into the corresponding boxed structures in the sgRNA1 promoter. The mutant names are listed above the boxes. In 1SL2, helices are designated as 1A, 1B, and 1D; the central region is designated as 1C; and the apical loop is designated as 1E. The right-angled arrow indicates the transcription start site (nt 2178) of sgRNA1. (B and C) Ectopic sgRNA3 syntheses of the mutants were analyzed by Northern blotting. Positions of gRNA, sgRNA1, sgRNA2, and sgRNA3 are indicated.

envisaged as similar large stem-loops (1SL2 and 2SL2) just upstream of the transcription start site adjacent to a smaller stem-loop (1SL1 and 2SL1) further upstream of the same site, indicating that both sgRNA1 and sgRNA2 promoters shared similar secondary structures.

Analysis of essential primary sequence and secondary structures for sgRNA1 promoter. To elucidate the involvement of

primary sequence and secondary structural elements in sgRNA1 synthesis, the two stem-loops (1SL1 and 1SL2) of the putative sgRNA1 promoter were analyzed by deletion or site-directed mutagenesis. The total RNA isolated from kenaf protoplasts 24 hpi was analyzed by Northern blot hybridization.

The potential structure of the sgRNA1 promoter sequence contained single-stranded regions at the 5′ and 3′ termini. The
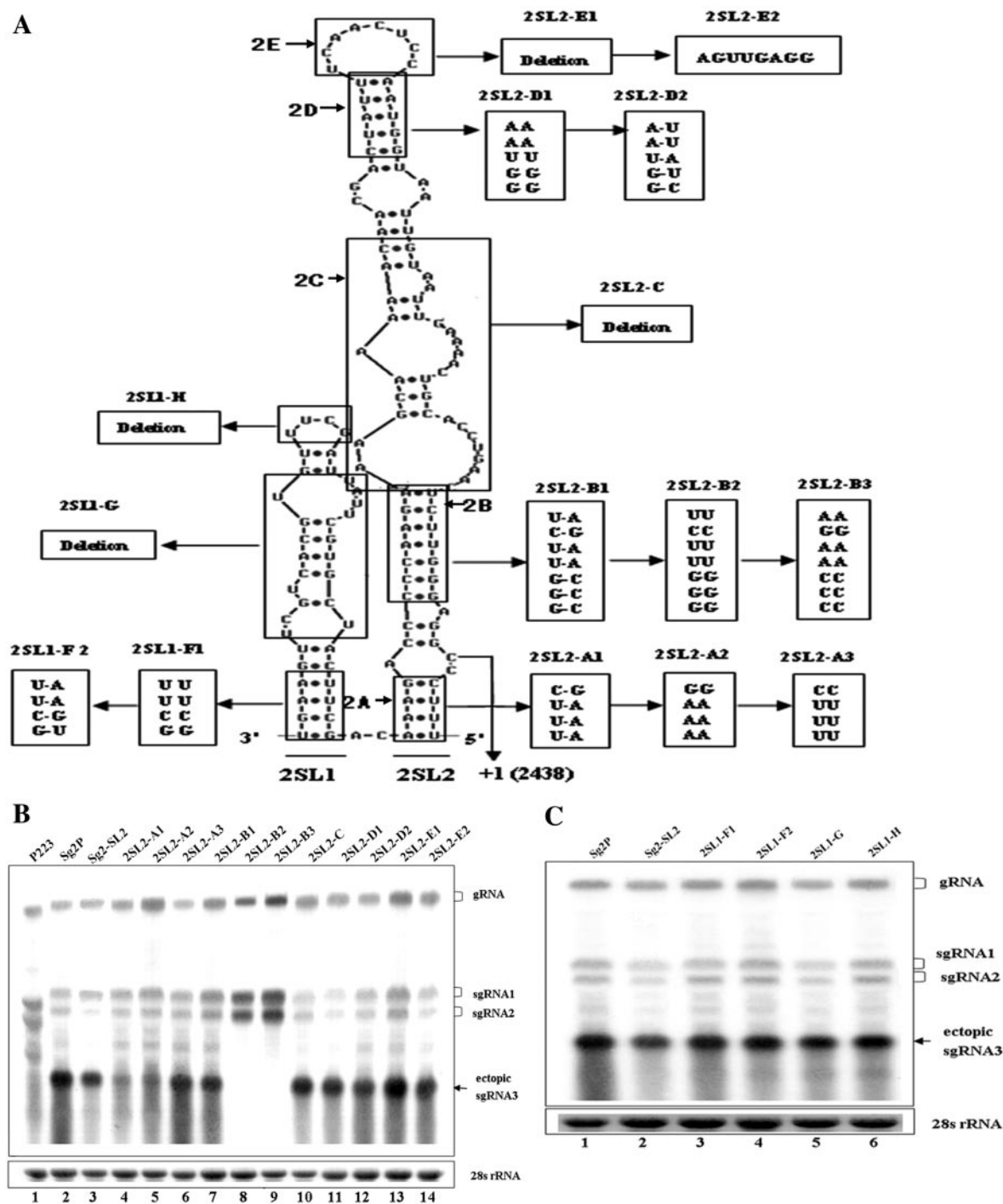
FIG. 7. Analyses of RNA sequence and secondary structure of sgRNA2 promoter. (A) Mutations in the 2SL1 and 2SL2 regions. The sequence is in the minus strand. In 2SL2, helices are designated as 2A, 2B, and 2D; the central region is designated as 2C; and the apical loop is designated as 2E. The transcription start site (nt 2438) of sgRNA2 is indicated with a right-angled arrow. (B and C) Northern blot analyses of the activity of sgRNA2 promoter mutants. All designations and methods are as described for Fig. 6. Positions of gRNA, sgRNA1, sgRNA2, and sgRNA3 are indicated.

5′ end of the single-stranded 3′-CUUU-5′ is important for sgRNA1 promoter activity (Fig. 4C, lane 6). To examine the role of the single-stranded RNA at the 3′ terminus, the region from nt 2052 to 2066 was deleted, and the promoter activity in the resultant mutant, Sg1S (Fig. 6B, lane 4), exhibited the same

level as that of the duplicated putative full-length sgRNA1 promoter by mutant Sg1P (Fig. 6B, lane 1), indicating that this sequence (nt 2052 to 2066) is redundant for sgRNA1 promoter activity. Thus, this promoter sequence is redefined as the complementary sequence of nt 2067 to 2183.

To analyze the role of stem-loop structure 1SL2 in sgRNA1 promoter activity, the boxed sections 1A, 1B, 1C, 1D, and 1E as well as the bulge between regions 1C and 1D were mutated (Fig. 6A). Disruption of the base pairing in section 1A did not affect the production of ectopic sgRNA3 in infected protoplasts (mutant 1SL2-A1; Fig. 6B, lane 5). However, reversion of the upper and lower base pairs abolished promoter activity (mutant 1SL2-A2; Fig. 6B, lane 6). Together, the two mutants indicated that the U residue and C residue flanking the transcription start site at positions 2177 and 2179 were important for promoter activity. Disruption of base pairing in section 1B reduced promoter activity to 20% of the wt level (mutant 1SL2-B1; Fig. 6B, lane 7), whereas restoration of base pairing by a compensatory mutation, 1SL2-B2, resulted in an increase in promoter activity to 67% of the wt level (Fig. 6B, lane 8). This indicated that, in contrast to section 1A, base pairing in section 1B was essential for promoter activity.

To test the role of the upper helix 1D of the 1SL2, nucleotide alterations that disrupted and restored the base pairing were introduced into mutants 1SL2-D1 and 1SL2-D2, respectively. The sgRNA promoter activity derived from the two mutants was comparable to the wt level (103% and 108%; Fig. 6B, lanes 11 and 12). As for the apical loop 1E of 1SL2, both mutant 1SL2-E1 and mutant 1SL2-E2, with deletion and replacement of the sequence (in the minus strand) 3′-AAAUA C-5′ with 3′-UUUAUG-5′, respectively, showed similar levels (91% to 99%) of promoter activity (Fig. 6B, lanes 13 and 14). For the central region 1C, complete nucleotide deletion was carried out to generate mutant 1SL2-C1. Surprisingly, the accumulation of sgRNA3 in the mutant 1SL2-C1 was increased to 109% of the wt level (Fig. 6B, lane 9). The deletion mutant 1SL2-C2 for the bulge (3′-ACU-5′) generated only 78% of the wt level (Fig. 6B, lane 10). Taken together, the bulge and the apical loop 1E were of minor importance for the full activity of the sgRNA1 promoter. The central region 1C and the upper helix 1D seemed to be dispensable for promoter activity.

To determine the possible role of the putative 1SL1, it was completely deleted as mutant sg1-SL2, which resulted in reduction of sgRNA3 to 60% of the wt level (Fig. 6C, lane 2). We then investigated the contributions of different elements in 1SL1 by a set of mutants, 1SL1-F1, 1SL1-F2, 1SL2-F3, 1SL1-G, 1SL1-H1, and 1SL1-H2 (Fig. 6A). The sgRNA3 levels of the mutants were slightly lower than the wt level (Fig. 6C, lanes 3 to 8). These data suggest that the whole 1SL1 is of minor importance for maintaining the level of sgRNA1.

Analysis of essential primary sequence and secondary structures for sgRNA2 promoter. We next characterized the primary sequence and secondary structural elements of 2SL1 and 2SL2 in the sgRNA2 promoter, respectively. Since the sgRNA2 promoter was able to generate ectopic sgRNAs similarly to the sgRNA1 promoter, the same strategy applied for sgRNA1 promoter analysis was used.

To test the significance of stem-loop 2SL2 in sgRNA2 promoter activity, five boxed sections, 2A, 2B, 2C, 2D, and 2E, were mutated (Fig. 7A). For section 2A, disruption of base pairing in mutant 2SL2-A3 had no effect on the transcription of sgRNA3 (Fig. 7B, lane 6). However, mutant 2SL2-A2 and the compensatory mutant 2SL2-A1 all exhibited relatively low transcription levels (35%) of sgRNA3 (Fig. 7B, lanes 4 and 5), indicating that the nucleotide sequence (3′-CUUU-5′) on the

right arm of section 2A, and not the secondary structure, was essential for promoter activity. This result is consistent with that obtained from the study with direct sequence deletion of 3′-CUUU-5′ in the putative sgRNA2 promoter (Fig. 5C, lane 5). Disruption of base pairing in section 2B resulted in no detectable ectopic sgRNA3 (mutants 2SL2-B2 and 2SL2-B3; Fig. 7B, lanes 8 and 9), whereas by restoration of the base pairing in a compensatory mutant, the promoter activity was restored to 67% of the wt level (2SL2-B1; Fig. 7B, lane 7). These results indicated that the secondary structure, but not the nucleotide sequence in section 2B, was absolutely indispensable for sgRNA2 transcription.

To determine the role of the upper portion of 2SL2 comprising the upper helix 2D, the same strategy as described above was employed to create mutants 2SL2-D1 and 2SL2-D2. Whether the helix 2D was disrupted (mutant 2SL2-D1) or restored (mutant 2SL2-D2), the transcription level of sgRNA3 (80% to 85% of the wt level) was not obviously different (Fig. 7B, lanes 11 and 12). As for the apical loop of 2SL2, mutants 2SL2-E1 and 2SL2-E2, which had deletion of 3′-UCAACUC C-5′ or replacement with 3′-AGUUGAGG-5′ in the minus strand, reduced promoter activity to 74% and 78% of the wt level, respectively (Fig. 7B, lanes 13 and 14). For the central region 2C, deletion mutant 2SL2-C produced 71% of the wt level (Fig. 7B, lane 10). Taken together, the central region 2C, the upper helix 2D, and the apical loop 2L were of minor importance for the full promoter activity of sgRNA2.

To investigate the role of 2SL1, mutant sg2-SL2 was created by deletion of the entire 2SL1, and the sgRNA3 accumulation was reduced to 76% of the wt level (Fig. 7C, lane 2). To further dissect different parts in 2SL1, a series of mutants, 2SL1-F1, 2SL1-F2, 2SL1-G1, and 2SL1-H, was created (Fig. 7A). The promoter activity of each mutant was comparable to the wt level (Fig. 7C, lanes 3 to 6). It was determined from this approach that the 2SL1 of sgRNA2 promoter enhanced transcription, in a fashion similar to that of the 1SL1 in sgRNA1 promoter.

Defining the minimal sequences required for promoter activity of two subgenomic RNAs. Both sgRNA1 and sgRNA2 promoters shared similar secondary structures, including the helices (1A and 1B versus 2A and 2B) at their basal portions. To determine the minimal sequences required for promoter activity, two mutants, mini-sg1 and mini-sg2, were generated. Structural analysis revealed that the putative secondary structures of the two minimal promoters, containing stable stems rich in G/C pairing, were consistent with their native structures of 1SL2 and 2SL2 in the two promoters (Fig. 8A). Results from transfection of kenaf protoplasts with the transcripts of the two mutants indicated that the two minimal sequences were able to produce 16% and 9% of the wt transcription level for mini-sg1 and mini-sg2, respectively (Fig. 8B). These data further confirmed that the basal portion sequences (Fig. 6A and 7A, sequences extending immediately below the 1C or 2C box to helix 1A or 2A box, except that 1SL1 includes a tetranucleotide sequence of 5′-UUUC-3′) were essential for activity of the two sgRNA promoters.

The 5′ UTR and the last 87 nt at the 3′ UTR of HCRSV can function as subgenomic promoters in vivo. The possibility that
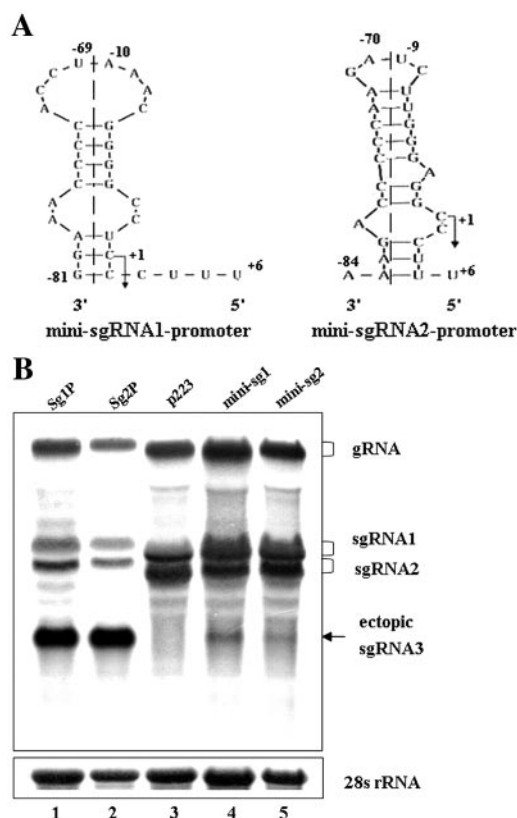
FIG. 8. Sequence and structure-specific requirements of sgRNA1 and sgRNA2 promoters. (A) RNA sequences and secondary structures of the minimal promoters for sgRNA1 and sgRNA2. The right-angled arrow represents the transcription start site (position +1). The numbers, relative to the native transcription start site, indicate the original positions of nucleotides within the promoter regions. The dashed line separates the promoter into two portions, which indicate that the minimal promoter is composed with two distal sequences of the original promoter. (B) Northern blot analysis of the activity of mini-sg1 and mini-sg2 promoter. Positions of gRNA, sgRNA1, sgRNA2, and sgRNA3 are indicated.

the minus-strand promoter of HCRSV can function as an sgRNA promoter was examined. Although the minus-strand promoter has not been mapped, it was previously shown that the last 87 nt at the 3′ UTR of HCRSV is essential for minus-strand synthesis (23). Thus, it is reasonable to predict that the 87-nt sequence may also function as a minus-strand promoter. The 5′ UTR of HCRSV, a 30-nt sequence, is further assumed to function as a plus-strand promoter. To determine the role of these two sequences as a possible sgRNA promoter, the 5′ UTR and the 87-nt sequence were duplicated and inserted into p223BS in forward or reverse orientation, respectively (Fig. 9A). Northern blot analysis showed that both sequences could generate ectopic sgRNA3, while their levels were lower than those of the two sgRNA promoters (Fig. 9B, lanes p5′UTR and p87). In addition, the ectopic sgRNA3 produced by the 87-nt sequence was smaller than the expected size (Fig. 9B, lane 5). In conclusion, this study demonstrated that both the 5′ UTR and the last 87 nt at the 3′ UTR of HCRSV could function as sgRNA promoters in vivo.
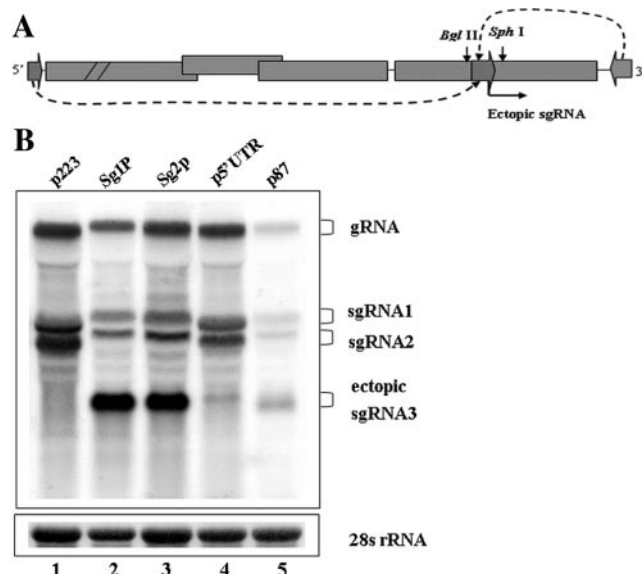

FIG. 9. The 5′ UTR and the last 87 nt at the 3′ UTR of HCRSV function as sgRNA promoters. (A) Diagram of constructs in which the sequence of the 5′ UTR (nt 1 to 30) or the last 87 nt of the 3′ UTR (nt 3825 to 3911) was duplicated and inserted into the BglII and SphI sites in the CP ORF of the p223BS mutant. The orientation of the inserted sequence is shown. (B) Northern blot analysis of sgRNA accumulation in protoplasts (24 hpi) transfected with in vitro transcripts of p5′UTR and p3′UTR. Bands corresponding to the ectopic sgRNA3 are indicated.

DISCUSSION

Among viruses with multiple sgRNAs, the sgRNA promoters often present great variations in their sizes, locations, and primary and secondary structures (17). BYDV has three sgRNA promoters, each with very different primary and secondary structures and locations relative to the transcription initiation sites (12, 13). Similarly, *Barley stripe mosaic virus* also has three different sgRNA promoters (9). The MP and CP sgRNA promoters of TMV have different primary and secondary structures (5). Conversely, both the TCV sgRNA promoters share similar positions and secondary structures. The promoter sequences are mostly located upstream of their transcription start sites and form extensive hairpin structures (24, 25).

During HCRSV replication, two 3′-coterminal sgRNAs are generated in infected kenaf protoplasts. Each sgRNA employs the guanine nucleotides located at nt 2178 and 2438 as its transcription start sites. When the G residue of the transcription start site was mutated, the corresponding sgRNA production was completely abolished. This agree with the observation that carmoviruses use internal initiation to generate sgRNAs (24). In most plant viruses, including BMV (4), TMV (5), BYDV (13), *Citrus leaf blotch virus* (22), and TCV (24), a guanine residue is at the transcription start site.

The sgRNA1 promoter of HCRSV was mapped to complementary nt 2067 to 2183, which overlapped with the 3′ end of ORF p81, while the sgRNA2 promoter extended from complementary nt 2312 to 2443, located in ORFs p8 and p9. With the exception of size, the two sgRNA promoters shared three common features in location and secondary structure. Both

were mostly located upstream of the transcription start site, with only 5 nt extended downstream of their start sites. Moreover, both sgRNA promoters possessed two stem-loops with similar secondary structures (1SL1 versus 2SL1 and 1SL2 versus 2SL2), although the sequence homology of the two promoters was only 36%. The large stem-loop structure (1SL2 versus 2SL2) of both sgRNA promoters could be divided into several regions (Fig. 6A and 7A), including the basal portion (5′-UUUC-3′, helices 1A and 1B versus 2A and 2B), central region (1C versus 2C), helices (1D versus 2D), and apical loop (1E versus 2E). Finally, both promoters had their essential elements at the basal portions, containing a conserved hexanucleotide (3′-CCCUUU-5′) and a stable stem structure rich in G/C pairing (Fig. 6A and 7A). A G/C-rich stem is also found in the basal portions of TCV sgRNA promoters (24, 25).

The minimal promoter sequences determined in the basal portions of both sgRNA promoters still showed promoter activity, although they are weaker than the wt. Computer analysis revealed that the secondary structures of the two minimal promoters were similar to the native structures of the sgRNA promoters. These results further confirmed that the hexanucleotide and the G/C-rich stem are essential for the sgRNA promoter activity of HCRSV. This is consistent with the well-documented evidence that sgRNA promoter activity requires both primary sequence and secondary structure (5, 12, 25). From the study of BMV and *Cowpea chlorotic mottle virus* sgRNA promoters, a sequence-specific recognition model was proposed (1, 2, 19). However, an alternate hairpin loop structure model was advocated for both AMV and BMV sgRNA promoters (6). Recent research revealed that a specific sequence in the BMV sgRNA promoter could direct replicase recognition and that the formation of a stem-loop is required at a step after replicase binding (20).

We questioned if the minus-strand gRNA promoter was active as an sgRNA promoter. The possibility had been demonstrated with two *Bromoviridae* viruses, BMV and AMV. With BMV, the minus-strand core promoter forms a stem-loop which shares an identical AUA triloop with the sgRNA promoter hairpin. In vitro assays showed that these two elements could be interchanged to a certain extent (7). Replacing the sgRNA promoter hairpin with the stem-loop results in more abundant sgRNA than that of the wt in vivo (20). For AMV, a triloop hairpin in the minus-strand promoter resembles the structure of its sgRNA promoter, and it can replace the authentic sgRNA promoter in live virus, but the sgRNA promoter fails to function as a minus-strand promoter in vivo (18). In our study, we have shown that both the 5′ UTR and the last 87-nt sequence at the 3′ UTR were able to produce sgRNAs, indicating that both genomic plus-strand and minus-strand promoters of HCRSV can function as sgRNA promoters in vivo. With the exception of the *Bromoviridae*, there is no other report on a minus-strand promoter serving as an sgRNA promoter and no plus-strand promoter functioning as an sgRNA promoter has been reported. HCRSV is the first virus identified with homology in function between the two sgRNA promoters and the putative genomic plus-strand and minus-strand promoters. There are host factors involved in viral replication and transcription (14). In HCRSV, the gRNA putative genomic plus-strand and minus-strand promoters were able to function as sgRNA promoters, suggesting that gRNA and sgRNA promoters may interact with some of the same host factors.

The evolution of sgRNA promoters is a very intriguing process. It has been hypothesized that sgRNA promoters evolved independently at the appropriate genome locations, while allowing overlapping ORFs to maintain their functions (12). In HCRSV, the two sgRNA promoters shared rather similar structures and two essential elements including a hexanucleotide sequence and a G/C-rich stem. It is possible that the sgRNA promoters may arise from a common origin, which develops via recombination or duplication and overlaps with ORFs during virus evolution.

REFERENCES

1. **Adkins, S., and C. C. Kao.** 1998. Subgenomic RNA promoters dictate the mode of recognition by bromoviral RNA-dependent RNA polymerases. Virology **252:**1–8.
2. **Adkins, S., R. W. Siegel, J. H. Sun, and C. C. Kao.** 1997. Minimal templates directing accurate initiation of subgenomic RNA synthesis in vitro by the brome mosaic virus RNA-dependent RNA polymerase. RNA **3:**634–647.
3. **Balmori, E., D. Gilmer, K. Richards, H. Guilley, and G. Jonard.** 1993. Mapping the promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3. Biochimie **75:**517–521.
4. **French, R., and P. Ahlquist.** 1988. Characterization and engineering of sequences controlling in vivo synthesis of brome mosaic virus subgenomic RNA. J. Virol. **62:**2411–2420.
5. **Grdzelishvili, V. Z., S. N. Chapman, W. O. Dawson, and D. J. Lewandowski.** 2000. Mapping of the tobacco mosaic virus movement protein and coat protein subgenomic RNA promoters in vivo. Virology **275:**177–192.
6. **Haasnoot, P. C., F. T. Brederode, R. C. Olsthoorn, and J. F. Bol.** 2000. A conserved hairpin structure in alfamovirus and bromovirus subgenomic promoters is required for efficient RNA synthesis in vitro. RNA **6:**708–716.
7. **Haasnoot, P. C., R. C. Olsthoorn, and J. F. Bol.** 2002. The brome mosaic virus subgenomic promoter hairpin is structurally similar to the iron-responsive element and functionally equivalent to the (−)-strand core promoter stem-loop C. RNA **8:**110–122.
8. **Huang, M., D. C. Koh, L. J. Weng, M. L. Chang, Y. K. Yap, L. Zhang, and S. M. Wong.** 2000. Complete nucleotide sequence and genome organization of hibiscus chlorotic ringspot virus, a new member of the genus *Carmovirus*: evidence for the presence and expression of two novel open reading frames. J. Virol. **74:**3149–3155.
9. **Johnson, J. A., J. N. Bragg, D. M. Lawrence, and A. O. Jackson.** 2003. Sequence elements controlling expression of barley stripe mosaic virus subgenomic RNAs in vivo. Virology **313:**66–80.
10. **Johnston, J. C., and D. M. Rochon.** 1995. Deletion analysis of the promoter for the cucumber necrosis virus 0.9-kb subgenomic RNA. Virology **214:**100–109.
11. **Kim, K. H., and C. Hemenway.** 1997. Mutations that alter a conserved element upstream of the potato virus X triple block and coat protein genes affect subgenomic RNA accumulation. Virology **232:**187–197.
12. **Koev, G., B. R. Mohan, and W. A. Miller.** 1999. Primary and secondary structural elements required for synthesis of barley yellow dwarf virus subgenomic RNA1. J. Virol. **73:**2876–2885.
13. **Koev, G., and W. A. Miller.** 2000. A positive-strand RNA virus with three very different subgenomic RNA promoters. J. Virol. **74:**5988–5996.
14. **Lai, M. M.** 1998. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology **244:**1–12.
15. **Levis, R., S. Schlesinger, and H. V. Huang.** 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. J. Virol. **64:**1726–1733.
16. **Liang, X. Z., S. W. Ding, and S. M. Wong.** 2002. Development of a kenaf (*Hibiscus cannabinus* L.) protoplast system for replication study of hibiscus chlorotic ringspot virus. Plant Cell Rep. **20:**982–986.
17. **Miller, W. A., and G. Koev.** 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. Virology **273:**1–8.
18. **Olsthoorn, R. C., P. C. Haasnoot, and J. F. Bol.** 2004. Similarities and differences between the subgenomic and minus-strand promoters of an RNA plant virus. J. Virol. **78:**4048–4053.
19. **Siegel, R. W., S. Adkins, and C. C. Kao.** 1997. Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. Proc. Natl. Acad. Sci. USA **94:**11238–11243.

20. **Sivakumaran, K., S. K. Choi, M. Hema, and C. C. Kao.** 2004. Requirements for brome mosaic virus subgenomic RNA synthesis in vivo and replicase-core promoter interactions in vitro. J. Virol. **78:**6091–6101.

21. **van der Vossen, E. A., T. Notenboom, and J. F. Bol.** 1995. Characterization of sequences controlling the synthesis of alfalfa mosaic virus subgenomic RNA in vivo. Virology **212:**663–672.

22. **Vives, M. C., L. Galipienso, L. Navarro, P. Moreno, and J. Guerri.** 2002. Characterization of two kinds of subgenomic RNAs produced by citrus leaf blotch virus. Virology **295:**328–336.

23. **Wang, H. H., and S. M. Wong.** 2004. Significance of the 3′-terminal region in (−)-strand RNA synthesis of hibiscus chlorotic ringspot virus. J. Gen. Virol. **85:**1763–1776.

24. **Wang, J., and A. E. Simon.** 1997. Analysis of the two subgenomic RNA promoters for turnip crinkle virus in vivo and in vitro. Virology **232:**174–186.

25. **Wang, J., C. D. Carpenter, and A. E. Simon.** 1999. Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. Virology **253:**327–336.

26. **White, K. A.** 2002. The premature termination model: a possible third mechanism for subgenomic mRNA transcription in (+)-strand RNA viruses. Virology **304:**147–154.

27. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.